

A finite element method for nonlinear elliptic problems

Article (Published Version)

Lakkis, Omar and Pryer, Tristan (2013) A finite element method for nonlinear elliptic problems. SIAM Journal on Scientific Computing, 35 (4). A2025-A2045. ISSN 1064-8275

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/46359/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

A FINITE ELEMENT METHOD FOR NONLINEAR ELLIPTIC PROBLEMS*

OMAR LAKKIS[†] AND TRISTAN PRYER[‡]

Abstract. We present a Galerkin method with piecewise polynomial continuous elements for fully nonlinear elliptic equations. A key tool is the discretization proposed in Lakkis and Pryer, 2011, allowing us to work directly on the strong form of a linear PDE. An added benefit to making use of this discretization method is that a *recovered (finite element) Hessian* is a byproduct of the solution process. We build on the linear method and ultimately construct two different methodologies for the solution of second order fully nonlinear PDEs. Benchmark numerical results illustrate the convergence properties of the scheme for some test problems as well as the Monge–Ampère equation and the Pucci equation.

Key words. fully nonlinear PDE, finite element method, Monge–Ampère equation, Pucci’s equation, finite element Hessian, nonvariational finite element method

AMS subject classifications. 65N30, 35J96

DOI. 10.1137/120887655

1. Introduction. Fully nonlinear partial differential equations (PDEs) arise in many areas, including differential geometry (the Monge–Ampère equation), mass transportation (the Monge–Kantorovich problem), dynamic programming (the Bellman equation), and fluid dynamics (the geostrophic equations). The numerical approximation of the solutions of such equations is thus an important scientific task. There are at least three main difficulties apparent when attempting to derive numerical methods for fully nonlinear equations. The first is the strong nonlinearity on the highest order derivative which generally precludes a variational formulation. The second is that a fully nonlinear equation does not always admit a classical solution, even if the problem data is smooth, and the solution has to be sought in a generalized sense (e.g., viscosity solutions), which is bound to slow convergence rates. The third, a common problem in nonlinear solvers, is that the exact solution may not be unique and constraints such as convexity requirements must be included to ensure uniqueness.

Regardless of the problems, the *numerical approximation of fully nonlinear second order elliptic equations*, as described in Caffarelli and Cabré (1995), has been the object of considerable recent research, particularly for the case of Monge–Ampère, of which Oliker and Prussner (1988), Loeper and Rapetti (2005), Dean and Glowinski (2006), Feng and Neilan (2009b), Oberman (2008), Awanou (2010), Davydov and Saeed (2012), Brenner et al. (2011), and Froese (2011) are selected examples.

For more general classes of fully nonlinear equations some methods have been presented; most notably, at least from a theoretical view point, in Böhmer (2008) the author presents a C^1 finite element method that shows stability and consistency (hence convergence) of the scheme. This follows a classical “finite difference” approach

*Submitted to the journal’s Methods and Algorithms for Scientific Computing section August 9, 2012; accepted for publication (in revised form) May 20, 2013; published electronically August 6, 2013.

<http://www.siam.org/journals/sisc/35-4/88765.html>

[†]Department of Mathematics, University of Sussex, Brighton BN1 9RF, United Kingdom (o.lakkis@sussex.ac.uk, <http://www.maths.sussex.ac.uk/Staff/OL>).

[‡]School of Mathematics, Statistics and Actuarial Sciences, University of Kent, Canterbury CT2 7NF, United Kingdom (t.pryer@kent.ac.uk). This author’s work was partially supported by an EPSRC D.Phil. scholarship grant and EPSRC grant EP/H024018/1.

outlined by Stetter (1973), which requires a high degree of smoothness on the exact solution. From a practical point of view this approach presents difficulties, in that the C^1 finite elements are hard to design and complicated to implement. A useful overview of Berger–Bernstein splines in two spatial dimensions is a full implementation is given in Davydov and Saeed (2012). Similar difficulties are encountered in finite difference methods and the concept of *wide-stencil* appears to be useful, for example, that by Kuo and Trudinger (1992, 2005), Oberman (2008), and Froese (2011). In particular, in Oberman (2008) the author uses the monotonicity arguments of Barles and Souganidis (1991) to construct convergent difference schemes to viscosity solutions for Monge–Ampère-type problems.

In Feng and Neilan (2009a), (2009b) and Awanou (2010) the authors give a method in which they approximate the general second order fully nonlinear PDE by a sequence of fourth order quasi-linear PDEs. These are quasi-linear biharmonic equations which are discretized via mixed finite elements or using high-regularity elements such as splines. In fact for the Monge–Ampère equation, which admits two solutions, of which one is convex and the other concave, this method allows for the approximation of both solutions via the correct choice of a parameter. On the other hand, although computationally less expensive than C^1 finite elements (an alternative to mixed methods for solving the biharmonic problem), the mixed formulation still results in an extremely large algebraic system and the lack of maximum principle for general fourth order equations makes it hard to apply vanishing viscosity arguments to prove convergence. In Brenner et al. (2011), the authors propose a penalty method for the Monge–Ampère equation very much in the spirit of a discontinuous Galerkin penalty method. In Awanou (2011), the author uses a “pseudo time” approach to capture the solution of a fully nonlinear problem. The method can be viewed as an explicit temporal discretization of a fully nonlinear parabolic problem.

It is worth citing also a *least squares* approach described by Dean and Glowinski (2006). This method consists in minimizing the mean square of the residual, using a Lagrange multiplier method. Also here a fourth order elliptic term appears in the energy.

In this paper, we depart from the above proposed methods and explore a more direct approach by applying the *nonvariational finite element method* (NVFEM), introduced in Lakkis and Pryer (2011a), as a solver for the Newton iteration directly derived from the PDE. To be more specific, consider the model problem

$$(1.1) \quad \mathcal{N}[u] := F(D^2u) - f = 0$$

with homogeneous Dirichlet boundary conditions where $f : \Omega \rightarrow \mathbb{R}$ is a prescribed function and $F : \text{Sym}(\mathbb{R}^{d \times d}) \rightarrow \mathbb{R}$ is a real-valued algebraic function of symmetric matrices, which provides an elliptic operator in the sense of Caffarelli and Cabré (1995), as explained below in Definition 2.1. The method we propose consists in applying a Newton’s method, given below by (4.2) of the PDE (1.1), which results in a sequence of linear nonvariational elliptic PDEs that can be dealt with using the NVFEM proposed in Lakkis and Pryer (2011a). The results in this paper are computational, so despite not having a complete proof of convergence, we test our algorithm on various problems that are specifically constructed to be well posed. In particular, we test our method on the Monge–Ampère problem, which is the de facto benchmark for numerical methods of fully nonlinear elliptic equations. This is in spite of Monge–Ampère having an extra complication, which is conditional ellipticity (the operator is elliptic only if the function is convex). A crucial, empirically observed

feature of our method is that the convexity (or concavity) is automatically preserved if one uses \mathbb{P}^2 elements or higher. For \mathbb{P}^1 elements this is not true and the scheme must be stabilized by reinforcing convexity (or concavity) at each Newton step. This was achieved in Pryer (2010) using a semidefinite programming method. A rigorous error analysis of our method applied to the Monge–Ampère problem on 2 dimensional domains can be found in Neilan (2012), where convergence is proved for sufficiently smooth solutions.

We stress that our method is not limited to the approximation of the Monge–Ampère problem. The Monge–Ampère problem is the only example of a fully nonlinear problem which has a hidden variational structure. See Remark 2.1 for details.

The rest of this paper is set out as follows. In section 2 we introduce some notation and the model problem, discuss its ellipticity, and present Newton’s method, which yields a sequence of nonvariational linearized PDEs. In section 3 we review of the NVFEM proposed in Lakkis and Pryer (2011a) and apply it to discretize the nonvariational linearized PDEs in Newton’s method. In section 4 we numerically exhibit the performance of our discretization on unconditionally elliptic fully nonlinear PDEs, those that are elliptic and well posed without constraining our solution to a certain class of functions. In section 5 we turn to conditionally elliptic problems by dealing with the prime example of such problems, i.e., Monge–Ampère. We apply the discretization to the Monge–Ampère equation making use of the work of Aguilera and Morin (2009) to check that *finite element convexity* is preserved at each iteration. Finally, in section 6 we address the approximation of Pucci’s equation, which is another important example of fully nonlinear elliptic equation.

All the numerical experiments for this research were carried out using the DOLFIN interface for FEniCS (Logg and Wells, 2010) and making use of Gnuplot and ParaView for the graphics.

2. Notation.

2.1. Functional setup. Let $\Omega \subset \mathbb{R}^d$ be an open and bounded Lipschitz domain. We denote $L_2(\Omega)$ to be the space of square (Lebesgue) integrable functions on Ω together with its inner product $\langle v, w \rangle := \int_{\Omega} vw$ and norm $\|v\| := \|v\|_{L_2(\Omega)} = \langle v, v \rangle^{1/2}$. We denote by $\langle v | w \rangle$ the action of a distribution v on the function w .

We use the convention that the derivative Du of a function $u : \Omega \rightarrow \mathbb{R}$ is a row vector, while the gradient of u , ∇u is the derivative’s transpose (an element of \mathbb{R}^d , representing Du in the canonical basis). Hence

$$(2.1) \quad \nabla u = (Du)^{\top}.$$

For second derivatives, we follow the common innocuous abuse of notation whereby the Hessian of u is denoted as D^2u (instead of the more consistent $D\nabla u$) and is represented by a $d \times d$ matrix.

The standard Sobolev spaces are as described by Ciarlet (1978) and Evans (1998),

$$(2.2) \quad H^k(\Omega) := W_2^k(\Omega) = \left\{ \phi \in L_2(\Omega) : \sum_{|\alpha| \leq k} D^{\alpha} \phi \in L_2(\Omega) \right\},$$

$$(2.3) \quad H_0^1(\Omega) := \text{closure of } C_0^{\infty}(\Omega) \text{ in } H^1(\Omega),$$

where $\alpha = (\alpha_1, \dots, \alpha_d)$ is a multi-index, $|\alpha| = \sum_{i=1}^d \alpha_i$, and derivatives D^{α} are understood in a weak sense.

We consider the case when the model problem (1.1) is uniformly elliptic in the following sense.

DEFINITION 2.1 (ellipticity, Caffarelli and Cabré (1995)). *The operator $\mathcal{N}[\cdot]$ in problem (1.1) is called elliptic on $\mathcal{C} \subseteq \text{Sym}(\mathbb{R}^{d \times d})$ if and only if for each $\mathbf{M} \in \mathcal{C}$ there exist $\Lambda \geq \lambda > 0$ that may depend on \mathbf{M} such that*

$$(2.4) \quad \lambda \sup_{|\boldsymbol{\xi}|=1} |\mathbf{N}\boldsymbol{\xi}| \leq F(\mathbf{M} + \mathbf{N}) - F(\mathbf{M}) \leq \Lambda \sup_{|\boldsymbol{\xi}|=1} |\mathbf{N}\boldsymbol{\xi}| \quad \forall \mathbf{N} \in \text{Sym}(\mathbb{R}^{d \times d}).$$

If the largest possible set \mathcal{C} for which (2.4) is satisfied is a proper subset of $\text{Sym}(\mathbb{R}^{d \times d})$ we say that $\mathcal{N}[\cdot]$ is conditionally elliptic.

The operator $\mathcal{N}[\cdot]$ in problem (1.1) is called uniformly elliptic if and only if for some $\lambda, \Lambda > 0$, called ellipticity constants, we have

$$(2.5) \quad \lambda \sup_{|\boldsymbol{\xi}|=1} |\mathbf{N}\boldsymbol{\xi}| \leq F(\mathbf{M} + \mathbf{N}) - F(\mathbf{M}) \leq \Lambda \sup_{|\boldsymbol{\xi}|=1} |\mathbf{N}\boldsymbol{\xi}| \quad \forall \mathbf{N}, \mathbf{M} \in \text{Sym}(\mathbb{R}^{d \times d}).$$

If F is differentiable, (2.5) can be obtained from conditions on the derivative of F . A generic $\mathbf{M} \in \text{Sym}(\mathbb{R}^{d \times d})$ is written as

$$(2.6) \quad \mathbf{M} = \begin{bmatrix} m_{1,1} & \dots & m_{1,d} \\ \vdots & \ddots & \vdots \\ m_{d,1} & \dots & m_{d,d} \end{bmatrix},$$

so the derivative of F in the direction \mathbf{N} is given by

$$(2.7) \quad \text{D}F(\mathbf{M})\mathbf{N} = F'(\mathbf{M}):\mathbf{N},$$

where the *gradient matrix* $F'(\mathbf{M})$ is defined by

$$(2.8) \quad F'(\mathbf{M}) := \begin{bmatrix} \partial F(\mathbf{M})/\partial m_{1,1} & \dots & \partial F(\mathbf{M})/\partial m_{1,d} \\ \vdots & \ddots & \vdots \\ \partial F(\mathbf{M})/\partial m_{d,1} & \dots & \partial F(\mathbf{M})/\partial m_{d,d} \end{bmatrix}.$$

Suppose F is differentiable. Then (2.4) is satisfied if and only if for each $\mathbf{M} \in \mathcal{C}$ there exists $\mu > 0$ such that

$$(2.9) \quad \boldsymbol{\xi}^\top F'(\mathbf{M})\boldsymbol{\xi} \geq \mu |\boldsymbol{\xi}|^2 \quad \forall \boldsymbol{\xi} \in \mathbb{R}^d.$$

Furthermore $\mathcal{C} = \text{Sym}(\mathbb{R}^{d \times d})$ and μ is independent of \mathbf{M} if and only if (2.5) is satisfied.

Assumption 2.1 (smooth elliptic operator). In the remainder of this paper we shall assume that $\mathcal{N}[\cdot]$ is conditionally elliptic on \mathcal{C} and

$$(2.10) \quad F \in C^1(\mathcal{C}).$$

Unless otherwise stated we will also assume that $\mathcal{C} = \text{Sym}(\mathbb{R}^{d \times d})$.

2.2. Newton's method. The smoothness requirement in Assumption 2.1 allows us to apply Newton's method to solve (1.1).

Given the initial guess $u^0 \in C^2(\Omega)$ with $\text{D}^2 u^0 \in \mathcal{C}$ for each $n \in \mathbb{N}_0$, find $u^{n+1} \in C^2(\Omega)$ with $\text{D}^2 u^{n+1} \in \mathcal{C}$ such that

$$(2.11) \quad \text{D}\mathcal{N}[u^n](u^{n+1} - u^n) = -\mathcal{N}[u^n],$$

where $\mathcal{D}\mathcal{N}[u]$ indicates the (Fréchet) derivative, which is formally given by

$$\begin{aligned} \mathcal{D}\mathcal{N}[u]v &= \lim_{\epsilon \rightarrow 0} \frac{\mathcal{N}[u + \epsilon v] - \mathcal{N}[u]}{\epsilon} \\ (2.12) \quad &= \lim_{\epsilon \rightarrow 0} \frac{F(\mathcal{D}^2 u + \epsilon \mathcal{D}^2 v) - F(\mathcal{D}^2 u)}{\epsilon} \\ &= F'(\mathcal{D}^2 u) : \mathcal{D}^2 v \end{aligned}$$

for each $v \in \mathbf{C}^2(\Omega)$. Combining (2.11) and (2.12) then results in the following non-variational sequence of linear PDEs. Given u^0 for each $n \in \mathbb{N}_0$ find u^{n+1} such that

$$(2.13) \quad F'(\mathcal{D}^2 u^n) : \mathcal{D}^2 (u^{n+1} - u^n) = f - F(\mathcal{D}^2 u^n).$$

The PDE (2.13) comes naturally in a nonvariational form. If we attempted to rewrite into a variational form in order to apply a “standard” Galerkin method, we would introduce an advection term, i.e., for generic v, w

$$(2.14) \quad F'(\mathcal{D}^2 v) : \mathcal{D}^2 w = \operatorname{div} [F'(\mathcal{D}^2 v) \nabla w] - \operatorname{div} [F'(\mathcal{D}^2 v)] \nabla w,$$

where the matrix divergence is taken rowwise,

$$(2.15) \quad \operatorname{div} [F'(\mathcal{D}^2 v(\mathbf{x}))] := \left(\sum_{i=1}^d \frac{\partial}{\partial x_i} [F'_{i,1}(\mathcal{D}^2 v(\mathbf{x}))], \dots, \sum_{i=1}^d \frac{\partial}{\partial x_i} [F'_{i,d}(\mathcal{D}^2 v(\mathbf{x}))] \right),$$

and the chain rule provides us, for each $j = 1, \dots, d$, with

$$(2.16) \quad \sum_{i=1}^d \frac{\partial}{\partial x_i} [F'_{i,j}(\mathcal{D}^2 v(\mathbf{x}))] = \sum_{k,l=1}^d \partial_{k,l} F'_{i,j}(\mathcal{D}^2 v(\mathbf{x})) \partial_{ikl} v(\mathbf{x}).$$

This procedure is undesirable for many reasons. First, it requires F to be twice differentiable and it involves a third order derivative of the functions u^{n+1} and u^n appearing in (2.11). Moreover, the “variational” reformulation could very well result in the problem becoming advection dominated and unstable for conforming FEM, as was manifested in numerical examples for the linear equation (Lakkis and Pryer (2011a, section 4.2)). It is worth noting that an exceptional case when this problem does not occur is given by the Monge–Ampère operator, as explained in Remark 2.1. In order to avoid these problems in general, we here propose the use of the NVFEM described in section 3.

Remark 2.1 (variational nature of Monge–Ampère operators). Among fully non-linear equations, whose (Fréchet) derivative is generally not in divergence form, the Monge–Ampère equations are an exception as they do exhibit a variational structure due to the divergence-free property of the linearization coefficient matrix. Namely, consider the Monge–Ampère equation

$$(2.17) \quad \det \mathcal{D}^2 u = f.$$

Recalling the basic determinant’s derivative formula for generic matrices \mathbf{X} and \mathbf{Y} ,

$$(2.18) \quad \mathcal{D}[\det \mathbf{X}] \mathbf{Y} = \operatorname{Cof} \mathbf{X} : \mathbf{Y},$$

and formally speaking, the derivative of the Monge–Ampère operator at the point u in the direction v is given by $\text{Cof } D^2u : D^2v$. Furthermore, since for any twice-differentiable w we have

$$(2.19) \quad \text{div}[\text{Cof } D^2w] = \mathbf{0}^\top,$$

the divergence structure of Monge–Ampère is revealed by an application of the divergence-product rule

$$(2.20) \quad \text{Cof } D^2u : D^2v = \text{div}[\text{Cof } D^2u \nabla v] - \text{div}[\text{Cof } D^2u] \nabla v.$$

It follows that the linearization of Monge–Ampère can be written in variational form. In fact, the Monge–Ampère problem is the Euler–Lagrange equation of the “energy-like” functional

$$(2.21) \quad \mathcal{J}[u] = \int_{\Omega} \frac{u}{d+1} \det D^2u - f =: \int_{\Omega} L(u, D^2u).$$

To verify this claim note that for a general second order variational problem the Euler–Lagrange equations are

$$(2.22) \quad 0 = D^2 : [\partial_2 L(u, D^2u)] + \partial_1 L(u, D^2u),$$

where ∂_i denotes the derivative with respect to the i th argument and the D^2 : operator’s application on a generic matrix-valued field \mathbf{V} is

$$(2.23) \quad D^2 : \mathbf{V} := \sum_{i,j=1}^d \partial_{i,j} v_{i,j} = \text{div}[(\text{div } \mathbf{V})^\top].$$

Hence the Euler–Lagrange equations for (2.21) are

$$(2.24) \quad 0 = D^2 : \left[\frac{u}{d+1} \text{Cof } D^2u \right] + \frac{1}{d+1} \det D^2u - f.$$

Expanding the first term in (2.24) we have

$$(2.25) \quad \begin{aligned} D^2 : \left[\frac{u}{d+1} \text{Cof } D^2u \right] &= \frac{1}{d+1} (D^2u : \text{Cof } D^2u + u D^2 : [\text{Cof } D^2u] \\ &\quad + 2 \text{div}[\text{Cof } D^2u] \nabla u). \end{aligned}$$

For a generic matrix \mathbf{X} it holds that

$$(2.26) \quad \mathbf{X} : \text{Cof } \mathbf{X} = d \det \mathbf{X}.$$

In addition, the divergence-free property given in (2.19) implies that the last two terms in (2.25) vanish, hence

$$(2.27) \quad D^2 : \left(\frac{u}{d+1} \text{Cof } D^2u \right) = \frac{d}{d+1} \det D^2u.$$

Substituting (2.27) into (2.24) yields the Monge–Ampère equation.

Note the degenerate nature of Monge–Ampère as a variational problem, in that the Euler–Lagrange equation of a functional involving second derivatives is only second order, whereas a fourth order equation would usually be expected. Indeed the terms that vanish in (2.25) are the fourth and third order u -derivative terms. This very special structure of the Monge–Ampère operator allows us in principle to avoid the NVFEM, whose use we study in this paper, by using a traditional FEM to solve the Newton’s method linear step. It is this precise fact that pushed us to explore NVFEM beyond the realm of Monge–Ampère where the divergence form is not available.

3. The NVFEM. In Lakkis and Pryer (2011a) we proposed the NVFEM to approximate the solution of problems of the form (2.13). We review here the NVFEM and explain how to use it in combination with the Newton method to derive a practical Galerkin method for the numerical approximation of the solution of (1.1).

3.1. Distributional form of (2.13) and generalized Hessian. Given $\mathbf{A} \in L_\infty(\Omega)^{d \times d}$ and for each $\mathbf{x} \in \Omega$, let $\mathbf{A}(\mathbf{x}) \in \text{Sym}(\mathbb{R}^{d \times d})$, the space of bounded, symmetric, positive definite, $d \times d$ matrices and $f : \Omega \rightarrow \mathbb{R}$. The *Dirichlet linear nonvariational elliptic problem* associated with \mathbf{A} and f is

$$(3.1) \quad \mathbf{A} : \mathbf{D}^2 u = f \text{ and } u|_{\partial\Omega} = 0.$$

Testing this equation, and assuming $u \in H^2(\Omega) \cap H_0^1(\Omega)$ such that $\nabla u|_{\partial\Omega} \in L_2(\partial\Omega)$, we may write it as

$$(3.2) \quad \langle \mathbf{A} : \mathbf{D}^2 u, \phi \rangle = \langle f, \phi \rangle \quad \forall \phi \in C_0^\infty(\Omega).$$

To allow a Galerkin-type discretization of (3.2), we need to restrict the test functions ϕ to finite element function spaces that are generally *not subspaces* of $H^2(\Omega)$. So before restricting, we need to extend and we use a traditional distribution theory (or generalized functions) approach. Given a function $v \in H^2(\Omega)$, let $\mathbf{n} : \partial\Omega \rightarrow \mathbb{R}^d$ be the outward pointing normal of Ω ; then the Hessian of v , $\mathbf{D}^2 v$, satisfies the following identity:

$$(3.3) \quad \langle \mathbf{D}^2 v, \phi \rangle = - \int_{\Omega} \nabla v \otimes \nabla \phi + \int_{\partial\Omega} \nabla v \otimes \mathbf{n} \phi \quad \forall \phi \in H^1(\Omega),$$

where $\mathbf{a} \otimes \mathbf{b} := \mathbf{a} \mathbf{b}^\top$ for \mathbf{a}, \mathbf{b} column vectors in \mathbb{R}^d . If $v \in H^1(\Omega)$ with $\nabla v|_{\partial\Omega} \in H^{-1/2}(\partial\Omega)$ the right-hand side of (3.3) still makes sense and defines $\mathbf{D}^2 v$ as an element in the dual of $H^1(\Omega)$ via

$$(3.4) \quad \langle \mathbf{D}^2 v | \phi \rangle := - \int_{\Omega} \nabla v \otimes \nabla \phi + \int_{\partial\Omega} \nabla v \otimes \mathbf{n} \phi \quad \forall \phi \in H^1(\Omega),$$

where $\langle \cdot | \cdot \rangle$ denotes the duality action on $H^1(\Omega)$ from its dual. We call $\mathbf{D}^2 v$ the *generalized Hessian* of v , and assuming that the coefficient tensor \mathbf{A} is in $C^0(\Omega)^{d \times d}$, for the product with a distribution to make sense, we now seek $u \in H_0^1(\Omega)$ such that $\nabla u|_{\partial\Omega} \in H^{-1/2}(\partial\Omega)$ and whose generalized Hessian satisfies

$$(3.5) \quad \langle \mathbf{A} : \mathbf{D}^2 v | \phi \rangle = \langle f, \phi \rangle \quad \forall \phi \in H^1(\Omega).$$

3.2. Finite element discretization and finite element Hessian. We discretize (3.5) with a standard piecewise polynomial approximation for test and trial spaces for both the problem variable, U , and the auxiliary (mixed-type) variable, $\mathbf{H}[U]$. Let \mathcal{T} be a conforming, shape regular triangulation of Ω , namely, \mathcal{T} is a finite family of sets such that

1. $K \in \mathcal{T}$ implies K is an open simplex (segment for $d = 1$, triangle for $d = 2$, tetrahedron for $d = 3$);
2. for any $K, J \in \mathcal{T}$ we have that $\overline{K} \cap \overline{J}$ is a full subsimplex (i.e., it is either \emptyset , a vertex, an edge, a face, or the whole of \overline{K} and \overline{J}) of both \overline{K} and \overline{J} ; and
3. $\bigcup_{K \in \mathcal{T}} \overline{K} = \overline{\Omega}$.

We use the convention where $h : \Omega \rightarrow \mathbb{R}$ denotes the *mesh-size function* of \mathcal{T} , i.e.,

$$(3.6) \quad h(\mathbf{x}) := \max_{K \ni \mathbf{x}} h_K,$$

where h_K is the *diameter* of an element K . We introduce the *finite element spaces*

$$(3.7) \quad \mathbb{V} := \{ \Phi \in H^1(\Omega) : \Phi|_K \in \mathbb{P}^p \forall K \in \mathcal{T} \text{ and } \Phi \in C^0(\Omega) \},$$

$$(3.8) \quad \mathring{\mathbb{V}} := \mathbb{V} \cap H_0^1(\Omega),$$

where \mathbb{P}^k denotes the linear space of polynomials in d variables of degree no higher than a positive integer k . We consider $p \geq 1$ to be fixed and denote by $\dot{N} := \dim \mathring{\mathbb{V}}$ and $N := \dim \mathbb{V}$. The discretization of problem (3.2) is then to find $(U, \mathbf{H}[U]) \in \mathring{\mathbb{V}} \times \mathbb{V}^{d \times d}$ such that

$$(3.9) \quad \begin{aligned} \langle \mathbf{H}[U], \Phi \rangle &= - \int_{\Omega} \nabla U \otimes \nabla \Phi + \int_{\partial\Omega} \nabla U \otimes \mathbf{n} \Phi \quad \forall \Phi \in \mathbb{V}, \\ \langle \mathbf{A} : \mathbf{H}[U], \Psi \rangle &= \langle f, \Psi \rangle \quad \forall \Psi \in \mathring{\mathbb{V}}. \end{aligned}$$

For an algebraic formulation of (3.9) we refer the reader to Lakkis and Pryer (2011a, section 2). Note that this discretization can be interpreted as a mixed method whereby the first (matrix) equation defines the *finite element Hessian* and the second (scalar) equation approximates the original PDE (3.2).

3.3. Two discretization strategies of (1.1). The finite element Hessian allows us two discretization strategies. The first strategy, detailed in section 4, consists in applying Newton first to set up (2.13) and then using the NVFEM (3.9) to solve each step. A second strategy becomes possible upon noting that given $U \in \mathbb{V}$ the finite element Hessian $\mathbf{H}[U]$ is a regular function,¹ which the generalized Hessian D^2U might fail to be. This allows us to apply nonlinear functions such as F to $\mathbf{H}[U]$ and consider the following *fully nonlinear finite element method*:

$$(3.10) \quad \begin{aligned} \langle \mathbf{H}[U], \Phi \rangle &= - \int_{\Omega} \nabla U \otimes \nabla \Phi + \int_{\partial\Omega} \nabla U \otimes \mathbf{n} \Phi \quad \forall \Phi \in \mathbb{V}, \\ \langle F(\mathbf{H}[U]), \Psi \rangle &= \langle f, \Psi \rangle \quad \forall \Psi \in \mathring{\mathbb{V}}. \end{aligned}$$

Of course, in order to solve the second equation, a finite-dimensional Newton method may be necessary (but this strategy leaves the door open for other nonlinear solvers, e.g., fixed point iterations). A finite element code based on this idea will be tested in section 6 to solve the Pucci equation.

¹A generalized function v is a *regular function*, or just *regular*, if it can be represented by a Lebesgue measurable function $f \in L_1^{\text{loc}}$ such that $\langle v | \phi \rangle = \int_{\Omega} f \phi$ for all $\phi \in C_0^\infty(\Omega)$. We follow the customary and harmless abuse in identifying v with f .

In summary the finite element Hessian allows both paths in the following diagram:

$$(3.11) \quad \begin{array}{ccc} \text{fully nonlinear PDE (1.1)} & \xrightarrow{\text{Newton}} & \text{nonvariational linear PDE's (2.13)} \\ \downarrow \text{FNFEM} & & \downarrow \text{NVFEM} \\ \text{fully nonlinear FE discretization (3.10)} & \xrightarrow{\text{Newton}} & \begin{array}{c} \text{discrete linear 1} \\ \text{discrete linear 2} \end{array} \end{array}$$

Although the diagram in (3.11) does not generally commute, if the function F is algebraically accessible, then it is commutative. By “algebraically accessible” we mean a function that can be computed in a finite number of algebraic operations or inverses thereof. In this paper, we use only algebraically accessible nonlinearities, but in principle, assuming derivatives are available, our methods could be extended to algebraically inaccessible nonlinearities, such as Bellman’s (or Isaacs’s) operators involving optimums over infinite families, e.g.,

$$(3.12) \quad F(\mathbf{M}) := \sup_{\alpha \in \mathcal{A}} \mathbf{L}_\alpha : \mathbf{M} \quad (\text{or } F(\mathbf{M}) := \inf_{\alpha \in \mathcal{A}} \sup_{\beta \in \mathcal{B}} \mathbf{L}_{\alpha, \beta} : \mathbf{M}),$$

where $\{\mathbf{L}_\alpha : \alpha \in \mathcal{A}\}$ (or $\{\mathbf{L}_{\alpha, \beta} : (\alpha, \beta) \in \mathcal{A} \times \mathcal{B}\}$) is a family of elliptic operators.

Remark 3.1 (relation to the vanishing moment method of Feng and Neilan (2009b)). In Feng and Neilan (2009b) the authors propose a mixed finite element method for the approximation of quasi-linear biharmonic problem

$$(3.13) \quad -\epsilon \Delta^2 u + F(\mathbf{D}^2 u) = 0$$

with an appropriately imposed extra boundary condition since the problem is fourth order. Modulo these boundary conditions the method given in (3.10) can be viewed as the *formal* limit of the vanishing moment method when discretized with a mixed method.

4. The discretization of unconstrained fully nonlinear PDEs. In this section we detail the application of the method reviewed in section 3 to the fully nonlinear model (1.1). Many fully nonlinear elliptic PDEs must be constrained in order to admit a unique solution. For example, the Monge–Ampère–Dirichlet problem is elliptic and admits a unique solution in the cone of convex (or concave) functions when $f > 0$ (or $f < 0$, respectively). Before we turn our attention to the more complicated constrained PDEs in section 5 we illustrate the Newton–NVFEM method in the simplest light. In this section we study fully nonlinear PDEs which have no such constraint.

Assumption 4.1 (unconditionally elliptic linearisation). We assume in this section that the Newton-step linearization (2.13) is elliptic. For this assumption to hold, it is sufficient to assume uniform ellipticity, i.e., (2.9) with $\mathcal{C} = \text{Sym}(\mathbb{R}^{d \times d})$ and $\mu > 0$ independent of \mathbf{M} .

4.1. The Newton–NVFEM method. Suppose we are given a boundary value problem of the following form: find $u \in H^2(\Omega) \cap H_0^1(\Omega)$ such that

$$(4.1) \quad \mathcal{N}[u] = F(\mathbf{D}^2 u) - f = 0 \quad \text{in } \Omega,$$

which satisfies Assumption 4.1.

Upon applying Newton's method to approximate the solution of problem (4.1) we obtain a sequence of functions $(u^n)_{n \in \mathbb{N}_0}$ solving the following linear equations in nonvariational form:

$$(4.2) \quad \mathbf{N}(\mathbf{D}^2 u^n): \mathbf{D}^2 u^{n+1} = g(\mathbf{D}^2 u^n),$$

where

$$(4.3) \quad \mathbf{N}(\mathbf{X}) := F'(\mathbf{X}),$$

$$(4.4) \quad g(\mathbf{X}) := f - F(\mathbf{X}) + F'(\mathbf{X}):\mathbf{X}.$$

The nonlinear finite element method to approximate (4.2) is: given an initial guess $U^0 := \Pi_0 u^0$ for each $n \in \mathbb{N}_0$ find $(U^{n+1}, \mathbf{H}[U^{n+1}]) \in \mathring{\mathbb{V}} \times \mathbb{V}^{d \times d}$ such that

$$(4.5) \quad \begin{aligned} \langle \mathbf{H}[U^{n+1}], \Phi \rangle + \int_{\Omega} \nabla U^{n+1} \otimes \nabla \Phi - \int_{\partial\Omega} \nabla U^{n+1} \otimes \mathbf{n} \Phi &= \mathbf{0} \quad \forall \Phi \in \mathbb{V} \\ \text{and } \langle \mathbf{N}(\mathbf{H}[U^n]): \mathbf{H}[U^{n+1}], \Psi \rangle &= \langle g(\mathbf{H}[U^n]), \Psi \rangle \quad \forall \Psi \in \mathring{\mathbb{V}}. \end{aligned}$$

4.2. Numerical experiments: A simple example. In this section we detail numerical experiments aimed at demonstrating the application of (4.5) to a simple model problem.

Example 4.1 (a simple fully nonlinear PDE). The first example we consider is a fully nonlinear PDE with a very smooth nonlinearity. The problem is

$$(4.6) \quad \begin{aligned} \mathcal{N}[u] &:= \sin(\Delta u) + 2\Delta u - f = 0 \text{ in } \Omega, \\ u &= 0 \text{ on } \partial\Omega, \end{aligned}$$

which is specifically constructed to be uniformly elliptic. Indeed

$$(4.7) \quad F'(\mathbf{D}^2 u) = (\cos(\Delta u) + 2)\mathbf{I},$$

which is uniformly positive definite. The Newton linearization of the problem is as follows: given u^0 , for $n \in \mathbb{N}_0$ find u^{n+1} such that

$$(4.8) \quad (\cos(\Delta u^n) + 2)\mathbf{I}: \mathbf{D}^2(u^{n+1} - u^n) = f - \sin(\Delta u^n) - 2\Delta u^n;$$

our approximation scheme is nothing but 4.5 with

$$(4.9) \quad \mathbf{N}(\mathbf{X}) = (\cos(\text{trace } \mathbf{X}) + 2)\mathbf{I},$$

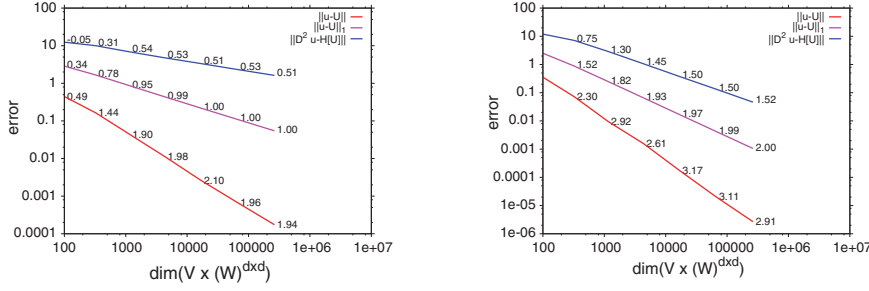
$$(4.10) \quad g(\mathbf{X}) = f - \sin(\text{trace } \mathbf{X}) - 2 \text{trace } \mathbf{X}.$$

Figure 4.1 details a numerical experiment on this problem when $d = 2$ and when $\Omega = [-1, 1]^2$ is a square which is triangulated using a criss-cross mesh.

Remark 4.1 (simplification of Example 4.1). Example 4.1 can be simplified considerably by noticing that

$$(4.11) \quad - \int_{\Omega} \text{tr } \mathbf{H}[U] \Phi = \int_{\Omega} (\nabla U)^{\top} \nabla \Phi \quad \forall \Phi \in \mathring{\mathbb{V}}.$$

This coincides with the definition of the *discrete Laplacian* and hence the NVFEM coincide with the standard conforming FEM, as described in Lakkis and Pryer (2011a,



(a) Taking \mathbb{V} to be the space of piecewise linear functions on Ω ($p = 1$). Notice that $\|u - U^M\| = O(h^2)$, $\|u - U^M\|_1 = O(h)$ and $\|D^2 u - \mathbf{H}[U^M]\| = O(h^{0.5})$.

(b) Taking \mathbb{V} to be the space of piecewise quadratic functions on Ω ($p = 2$). Notice that $\|u - U^M\| = O(h^3)$, $\|u - U^M\|_1 = O(h^2)$ and $\|D^2 u - \mathbf{H}[U^M]\| = O(h^{1.5})$.

FIG. 4.1. Numerical experiments for Example 4.1. Choose f appropriately such that $u(\mathbf{x}) = \sin(\pi x) \sin(\pi y)$. We use an initial guess $u^0 = 0$ and run the iterative procedure until $\|U^{n+1} - U^n\| \leq 10^{-8}$, setting $U := U^M$ as the final Newton iterate of the sequence. Here we are plotting log-log error plots together with experimental convergence rates for $L_2(\Omega)$, $H^1(\Omega)$ error functionals for the problem variable U and an $L_2(\Omega)$ error functional for the auxiliary variable, $\mathbf{H}[U]$.

Theorem 3.5). This observation applies to all fully nonlinear equations with nonlinearity of the form (1.1) with

$$(4.12) \quad F(\mathbf{M}) := a(\text{tr } \mathbf{M})$$

for some given a . This class of problems can be solved using a variational finite element method and can be used for comparison with our method.

Remark 4.2 (the method of Jensen and Smears (2011)). Jensen and Smears (2011) use a *modified* finite element method together with a *localization argument* in order to prove convergence of a finite element method for a class of Hamilton–Jacobi–Bellman equation, specifically, for those where the diffusion matrix is isotropic, i.e., $\mathbf{L}_\alpha = a_\alpha \mathbf{I}$ in (3.12), for some family $\{a_\alpha\}_\alpha$. The localization argument allows the use of an inconsistent projection operator, the Ritz projection, in the Barles and Souganidis (1991) convergence arguments.

While the methods do not coincide, their localization argument can be applied to our formulation for problems of the form (4.12) with an additional evolution term upon approximating the integrals appearing in (3.10) with a midpoint quadrature. We hope to extend the theory to encompass more general nonlinear operators using a *nonvariational Ritz projection*.

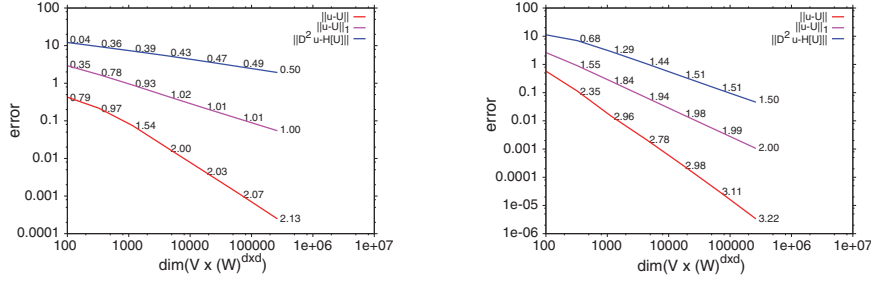
Example 4.2 (nonvariational example). This is a simple example where the *variational trick* mentioned in Remark 4.1 cannot be applied. We fix $d = 2$ and consider the problem

$$(4.13) \quad \begin{aligned} \mathcal{N}[u] &:= (\partial_{11}u)^3 + (\partial_{22}u)^3 + \partial_{11}u + \partial_{22}u - f = 0 \text{ in } \Omega, \\ u &= 0 \text{ on } \partial\Omega. \end{aligned}$$

The approximation scheme is then (4.5) with

$$(4.14) \quad \mathbf{N}(\mathbf{X}) := \begin{bmatrix} 3\mathbf{X}_{11}^2 + 1 & 0 \\ 0 & 3\mathbf{X}_{22}^2 + 1 \end{bmatrix},$$

$$(4.15) \quad g(\mathbf{X}) := f + 2(\mathbf{X}_{11}^3 + \mathbf{X}_{22}^3).$$



(a) Taking \mathbb{V} to be the space of piecewise linear functions on Ω ($p = 1$). Notice that $\|u - U^M\| = O(h^2)$, $\|u - U^M\|_1 = O(h)$ and $\|D^2 u - H[U^M]\| \approx O(h^{0.5})$.

(b) Taking \mathbb{V} to be the space of piecewise quadratic functions on Ω ($p = 2$). Notice that $\|u - U^M\| = O(h^3)$, $\|u - U^M\|_1 = O(h^2)$ and $\|D^2 u - H[U^M]\| = O(h^{1.5})$.

FIG. 4.2. Numerical experiments for Example 4.2. Choose f appropriately such that $u(\mathbf{x}) = \sin(\pi x) \sin(\pi y)$. We use an initial guess $u^0 = 0$ and run the iterative procedure until $\|U^{n+1} - U^n\| \leq 10^{-8}$, setting $U := U^M$ as the final Newton iterate of the sequence. Here we are plotting log-log error plots together with experimental convergence rates for $L_2(\Omega)$, $H^1(\Omega)$ error functionals for the problem variable U and an $L_2(\Omega)$ error functional for the auxiliary variable, $H[U]$.

Figure 4.2 details a numerical experiment on this problem in the case $d = 2$ and $\Omega = [-1, 1]^2$ is triangulated with a criss-cross mesh. A similar example is studied in Davydov and Saeed (2012, Example 5.2) using Böhmer's method.

5. The Monge–Ampère–Dirichlet problem. In this section we apply the Newton–NVFEM to the Monge–Ampère–Dirichlet (MAD) problem

$$(5.1) \quad \begin{aligned} \det D^2 u &= f \text{ in } \Omega, \\ u &= g \text{ on } \partial\Omega. \end{aligned}$$

Our numerical experiments exhibit the robustness of our method when computing (smooth) classical solutions of the MAD equation. Most importantly we noted the following facts:

- (i) the use of \mathbb{P}^p elements with $p \geq 2$ is essential as \mathbb{P}^1 do not work (at least on general meshes),
 - (ii) the convexity of the Newton iterates is conserved throughout the computation.
- Our observations are purely empirical from computations, which leaves it as an interesting open problem to prove this property.

Remark 5.1 (the MAD problem fails to satisfy Assumption 4.1). To clarify Assumption 4.1 for the MAD problem (5.1), in view of the characteristic expansion of determinant we have

$$(5.2) \quad F'(\mathbf{X}) = \text{Cof } \mathbf{X}.$$

This implies that the linearization of MAD is well posed only if we restrict the class of functions we consider to those u that satisfy

$$(5.3) \quad \boldsymbol{\xi}^\top \text{Cof } D^2 u \boldsymbol{\xi} \geq \lambda |\boldsymbol{\xi}|^2 \quad \forall \boldsymbol{\xi} \in \mathbb{R}^d$$

for some (u -dependent) $\lambda > 0$. Note that (5.3) is equivalent to the following two conditions as well:

$$(5.4) \quad \begin{aligned} \boldsymbol{\xi}^\top D^2 u \boldsymbol{\xi} &\geq \lambda |\boldsymbol{\xi}|^2 \quad \forall \boldsymbol{\xi} \in \mathbb{R}^d, \\ u &\text{ is strictly convex.} \end{aligned}$$

Loeper and Rapetti (2005) have shown that for the *continuous* (infinite-dimensional) Newton method described in 2.2, given a strictly convex initial guess u^0 , each iterate u^n will be convex. It is crucial that this property is preserved at the discrete level, as it guarantees the solvability of each iteration in the *discretized* Newton method. For this the right notion of convexity turns out to be the finite element convexity as developed in Aguilera and Morin (2009). In Pryer (2010), an intricate method based on semidefinite programming provided a way to constrain the solution in the case of \mathbb{P}^1 elements. Here we observe that the finite element convexity is automatically preserved, provided we use \mathbb{P}^2 or higher conforming elements.

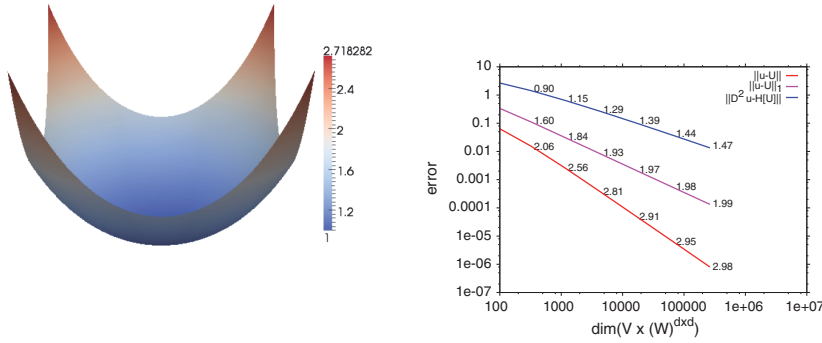
5.1. Newton's method applied to Monge–Ampère. In view of (5.2) it is clear that

$$(5.5) \quad D\mathcal{N}[u]v = \text{Cof } D^2u : D^2v.$$

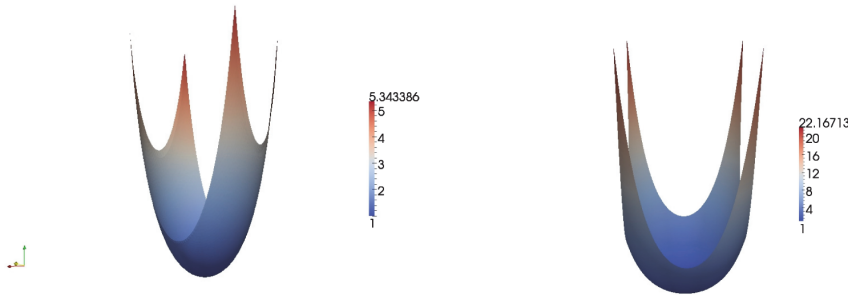
Applying the methodology set out in section 4 we set

$$(5.6) \quad \mathbf{N}(D^2u^n) = \text{Cof } D^2u^n \text{ and}$$

$$(5.7) \quad g(D^2u^n) = f - \det D^2u^n + \text{Cof } D^2u^n : D^2u^n.$$



(a) The FE approximation to the function $u(\mathbf{x}) = \exp\left(\frac{|\mathbf{x}|^2}{2}\right)$. (b) Log-log error plot for \mathbb{P}^2 Lagrange finite elements.



(c) The principal minor of $\mathbf{H}[U]$, an approximation to the Hessian of the function $u(\mathbf{x}) = \exp\left(\frac{|\mathbf{x}|^2}{2}\right)$.

(d) The determinant of $\mathbf{H}[U]$.

FIG. 5.1. Numerical results for the MAD problem on the square $S = [-1, 1]^2$. We choose the problem data f and g appropriately such that the solution is the radially symmetric function $u(\mathbf{x}) = \exp(|\mathbf{x}|^2/2)$. We plot the finite element solution together with a log-log error plot for various error functionals as in Figure 4.1.

5.2. Numerical experiments. In this section we study the numerical behavior of the scheme presented in Definition 4.1 applied to the MAD problem.

We present a set of benchmark problems constructed from the problem data such that the solution to the Monge–Ampère equation is known. We fix Ω to be the square $S = [-1, 1]^2$ or $[0, 1]^2$ (specified in the problem) and test convergence rates of the discrete solution to the exact solution.

Figures 5.1–5.3 detail the various experiments and show numerical convergence results for each of the problems studied as well as solution plots. It is worthy of note that each of the solutions seems to be convex; however, this is not necessarily the case. They are all, however, finite element convex Aguilera and Morin (2009). In each of these cases the Dirichlet boundary values are not zero. The implementation of nontrivial boundary conditions is described in Lakkis and Pryer (2011a, section 3.6) or in more detail in Pryer (2010, section 4.4).

Remark 5.2 (the initial guess to the Newton method). As with any Newton method we require a starting guess, not just for U^0 but also of $\mathbf{H}[U^0]$. Due to the mild nonlinearity with the previous example an initial guess of $U^0 \equiv 0$ and $\mathbf{H}[U^0] \equiv \mathbf{0}$ was sufficient. The initial guess to the MAD problem must be more carefully sought.

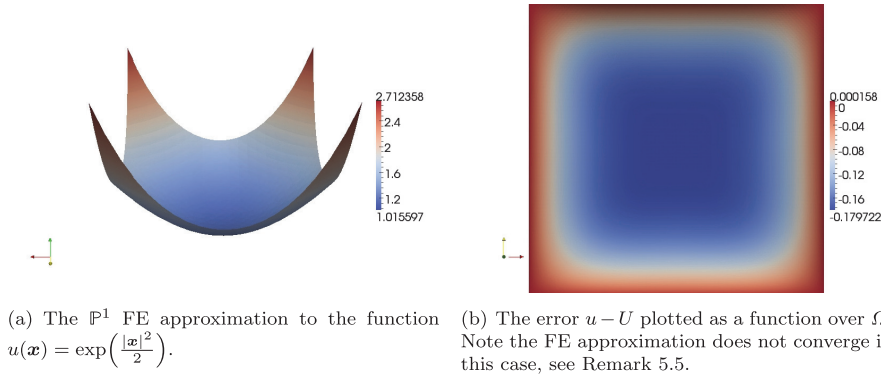


FIG. 5.2. Numerical results for the MAD problem on the square $S = [-1, 1]^2$. We choose the problem data f and g appropriately such that the solution is the radially symmetric function $u(\mathbf{x}) = \exp(|\mathbf{x}|^2/2)$.

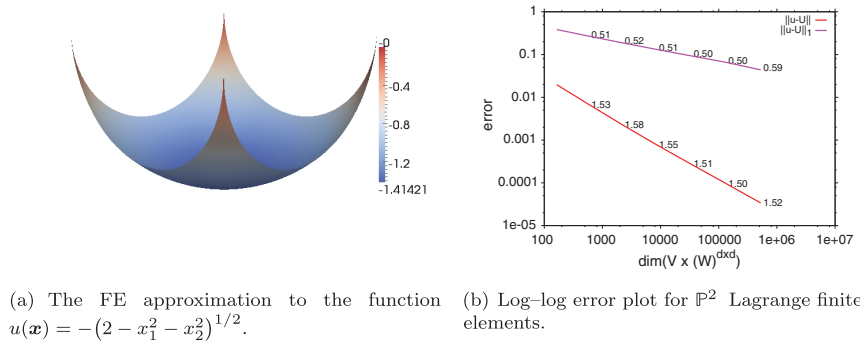


FIG. 5.3. Numerical results for the MAD problem on the square $S = [-1, 1]^2$. Choose f and g appropriately such that the solution is $u(\mathbf{x}) = -(2 - x_1^2 - x_2^2)^{1/2}$. Note the function has singular derivatives on the corners of S . We plot the finite element solution together with a log-log error plot for various error functionals as in Figure 4.1.

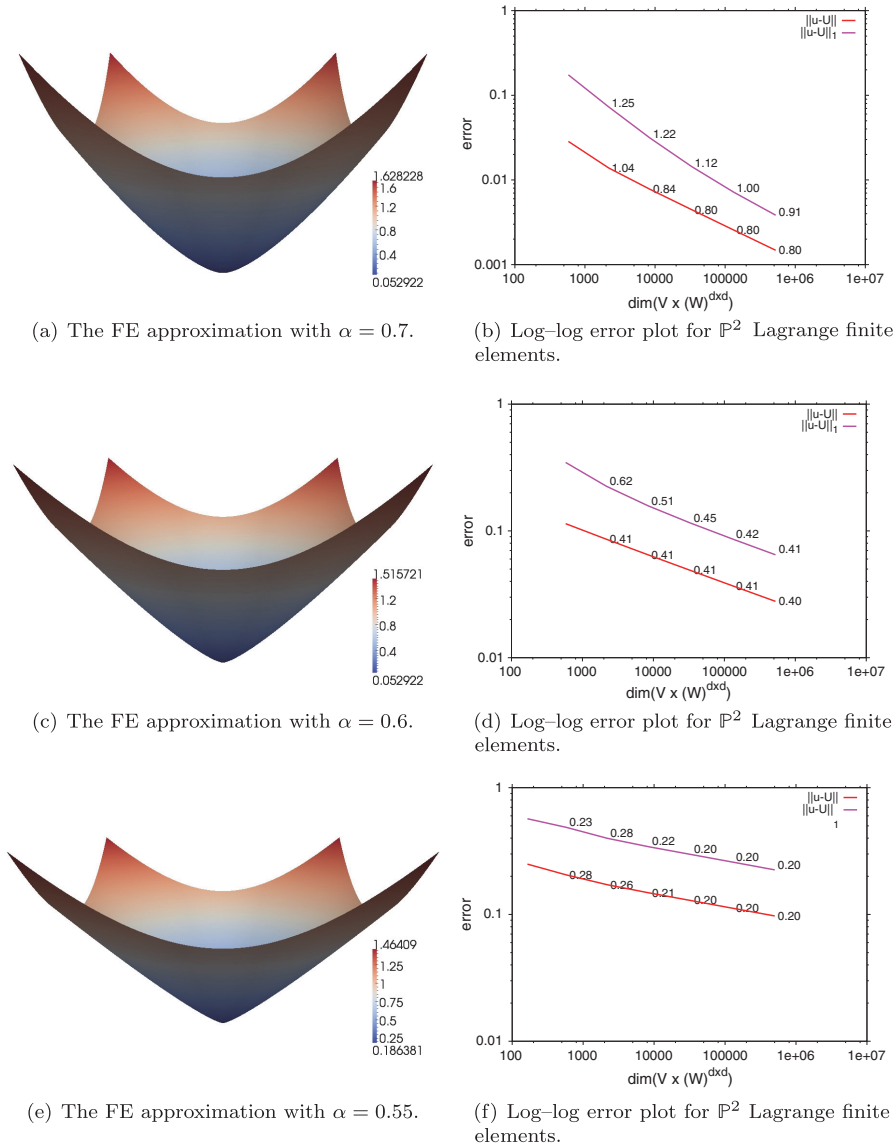


FIG. 5.4. Numerical results for the MAD problem on the square $S = [-1, 1]^2$. We choose f and g appropriately such that the solution is $u(\mathbf{x}) = |\mathbf{x}|^{2\alpha}$ with various α . Note the function has singular derivatives at the origin, and the mesh is irregular such that the singularity does not lie on a node. We plot the finite element solution together with a log-log error plot for various error functionals as in Figure 4.1.

Following a trick, described in Dean and Glowinski (2003), we chose $(U^0, \mathbf{H}[U^0])$ to be the nonvariational finite element approximation of u^0 such that

$$(5.8) \quad \Delta u^0 = 2\sqrt{f} \text{ in } \Omega,$$

$$(5.9) \quad u^0 = g \text{ on } \partial\Omega.$$

Remark 5.3 (degree of the FE space). In the previous example the lowest order convergent scheme was found by taking \mathbb{V} to be the space of piecewise linear functions

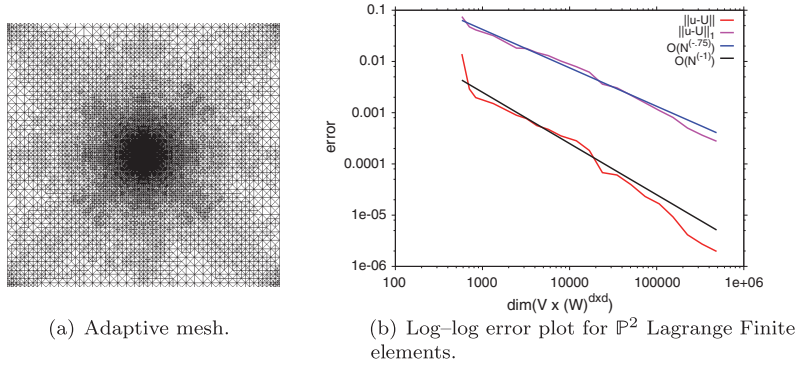


FIG. 5.5. Numerical results for a solution to the Monge–Ampère–Dirichlet equation with f and g appropriately such that the solution is $u(\mathbf{x}) = |\mathbf{x}|^{2\alpha}$ with $\alpha = 0.55$. We choose $p = 2$ and use an adaptive scheme based on ZZ gradient recovery. The mesh is refined correctly about the origin. Note that when $\dim \mathbb{V} = 20,420$ the adaptive solution achieves $\|u - U^M\| \approx 0.000067$, and the uniform solution given in Figure 5.4 satisfies $\|u - U^M\| \approx 0.18$ using the same number of degrees of freedom. Using the adaptive strategy $\|u - U^M\|$ converges like $O(N^{-1})$ and $|u - U^M|_1$ converges like $O(N^{-3/4})$.

($p = 1$). For the MAD problem we require a higher approximation power, hence we take \mathbb{V} to be the space of piecewise quadratic functions, i.e., $p = 2$.

Although the choice of $p = 1$ gives a stable scheme, convergence is not achieved. This can be characterized by Aguilera and Morin (2009, Theorem 3.6), which roughly says you require more approximation power than what piecewise linear functions provide to be able to approximate all convex functions. Compare with Figure 5.2.

5.3. Nonclassical solutions. The numerical examples given in Figures 5.1–5.3 both describe the numerical approximation of classical solutions to the MAD problem. In the case of Figure 5.1 $u \in C^\infty(\overline{\Omega})$, whereas in Figure 5.3 $u \in C^\infty(\Omega) \cap C^0(\overline{\Omega})$. We now take a moment to study less regular solutions, i.e., viscosity solutions which are not classical. In this test the solution is given by

$$(5.10) \quad u = |\mathbf{x}|^{2\alpha}$$

for $\alpha \in (1/2, 3/4)$. The exact solution $u \notin H^2(\Omega)$. In Figure 5.4 we discuss the convergence properties of the method with respect to α . Note that for $\alpha < 1/2$ since the function u is not convex, the Monge–Ampère operator is not elliptic. As a result the numerical scheme does not converge for $\alpha < 1/2$. Finally, in Figure 5.5 we conduct an adaptive experiment based on a gradient recovery a posteriori estimator. The recovery estimator we make use of is the Zienkiewicz–Zhu (ZZ) patch recovery technique; see Lakkis and Pryer (2011b), Pryer (2010, section 2.4), or Ainsworth and Oden (2000, section 4) for further details.

6. Pucci’s equation. In this section we look to discretize the nonlinear problem given by Pucci’s equation as a system of nonlinear equations. Pucci’s equation arises as a linear combination of Pucci’s extremal operators. It can nevertheless be written in an algebraically accessible form, without the need to compute the eigenvalues.

DEFINITION 6.1 (Pucci's extremal operators, Caffarelli and Cabré (1995)). Let $\mathbf{N} \in \text{Sym}(\mathbb{R}^{d \times d})$ and $\sigma(\mathbf{N})$ be its spectrum; then the extremal operators are

$$(6.1) \quad \mathcal{M}(\mathbf{N}) := \sum_{\lambda_i \in \sigma(\mathbf{N})} \alpha_i \lambda_i = 0$$

with $\alpha_i \in \mathbb{R}$. The maximal (minimal) operator, commonly denoted \mathcal{M}^+ (\mathcal{M}^-), has coefficients that satisfy

$$(6.2) \quad 0 < \alpha_1 \leq \dots \leq \alpha_n \quad (\alpha_1 \geq \dots \geq \alpha_n > 0),$$

respectively.

6.1. The planar case and uniform ellipticity. In the case $d = 2$ the normalized Pucci's equation reduces to finding u such that

$$(6.3) \quad \alpha \lambda_2 + \lambda_1 = 0,$$

where $\mathbf{N} := D^2 u$. Note that if $\alpha = 1$, (6.3) reduces to the Poisson–Dirichlet problem. This can be easily seen when reformulating the problem as a second order PDE (Dean and Glowinski (2005)). Making use of the characteristic polynomial, we see

$$(6.4) \quad \lambda_i = \frac{\Delta u + (-1)^i \left((\Delta u)^2 - 4 \det D^2 u \right)^{1/2}}{2}, \quad i = 1, 2.$$

Thus Pucci's equation can be written as

$$(6.5) \quad 0 = (\alpha + 1) \Delta u + (\alpha - 1) \left((\Delta u)^2 - 4 \det D^2 u \right)^{1/2},$$

which is a nonlinear combination of Monge–Ampère and Poisson problems. However, owing to the Laplacian terms, and unlike the Monge–Ampère–Dirichlet problem, Pucci's equation is (unconditionally) uniformly elliptic as

$$(6.6) \quad (\text{tr } \mathbf{X})^2 - 4 \det \mathbf{X} \geq 0 \quad \forall \mathbf{X} \in \mathbb{R}^{2 \times 2}.$$

The discrete problem we use is a direct approximation of (6.5); we seek $(U, \mathbf{H}[U])$ such that

$$(6.7) \quad \int_{\Omega} \left((\alpha + 1) \text{tr } \mathbf{H}[U] + (\alpha - 1) \left((\text{tr } \mathbf{H}[U])^2 - 4 \det \mathbf{H}[U] \right)^{1/2} \right) \Phi = 0,$$

$$(6.8) \quad \langle \mathbf{H}[U], \Psi \rangle = - \int_{\Omega} \nabla U \otimes \nabla \Psi + \int_{\partial \Omega} \nabla U \otimes \mathbf{n} \Psi \quad \forall (\Phi, \Psi) \in \mathring{\mathbb{V}} \times \mathbb{V}.$$

The result is a nonlinear system of equations which was solved using an algebraic Newton method.

6.2. Numerical experiments. We conduct numerical experiments to be compared with those of Dean and Glowinski (2005). The first problem we consider is a classical solution of Pucci's equation (6.3). Let $\mathbf{x} = (x, y)^T$; then the function

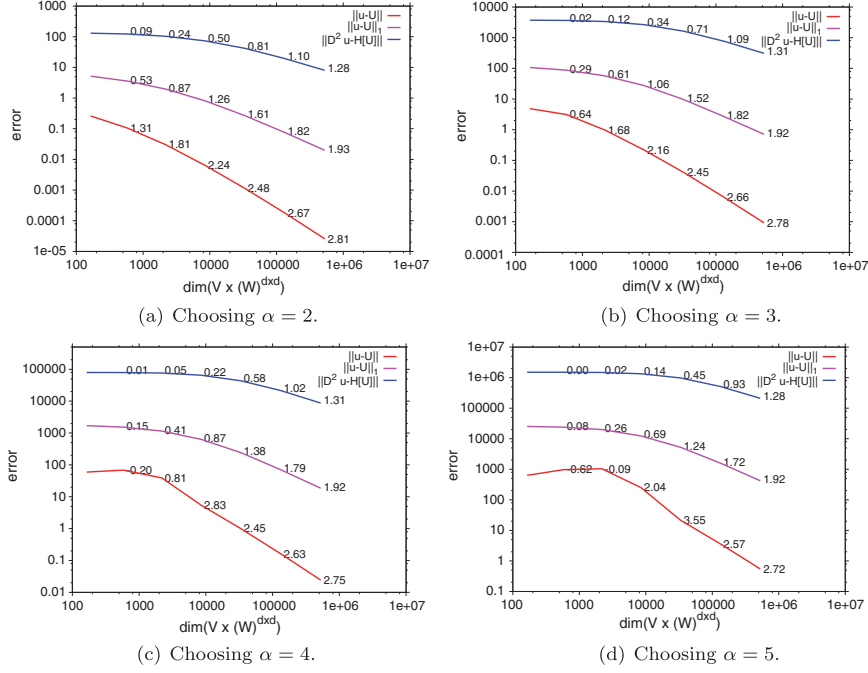


FIG. 6.1. Numerical results for a classical solution to Pucci's equation (6.9). As with the case of the MAD problem we choose $p = 2$. We use a Newton method to solve the algebraic system until the residual of the problem (see Kelley (1995)) is less than 10^{-10} (which is overkill to minimize Newton error effects). We plot log-log error plots with experimental orders of convergence for various norms and values of α .

$$(6.9) \quad u(\mathbf{x}) = -\left(\left((x+1)^2 + (y+1)^2\right)^{(1-\alpha)/2}\right)$$

solves Pucci's equation almost everywhere away from $(x, y) = (-1, -1)$ with $g := u|_{\partial\Omega}$. Let \mathcal{T} be an irregular triangulation of $\Omega = [-0.95, 1]^2$. In Figure 6.1 we detail a numerical experiment considering the case $\alpha \in [2, 5]$.

We also conduct a numerical experiment to be compared with Oberman (2008). In this problem we consider a solution of Pucci's equation with a piecewise defined boundary. Let $\Omega = [-1, 1]^2$ and the boundary data be given as

$$(6.10) \quad g(\mathbf{x}) := \begin{cases} 1 & \text{when } |x| \geq \frac{1}{2} \text{ and } |y| \geq \frac{1}{2}, \\ 0 & \text{otherwise.} \end{cases}$$

Figure 6.2 details the numerical experiment on this problem with various values of α .

Since the solution to the Pucci's equation with piecewise boundary (6.10) is clearly singular near the discontinuities, we have also conducted an adaptive experiment based on a gradient recovery a posteriori estimator (as in section 5.3). As can be seen from Figure 6.3, we regain qualitatively similar results using far fewer degrees of freedom.

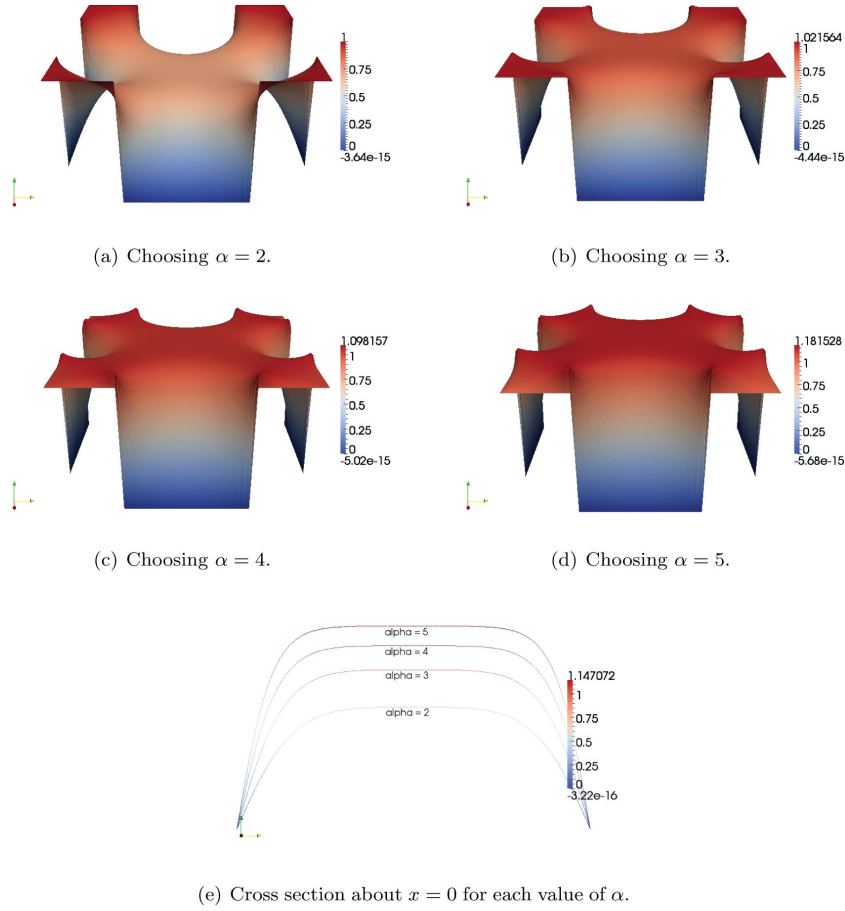


FIG. 6.2. Numerical results for a solution to Pucci's equation with a piecewise defined boundary condition (6.10). We choose $p = 2$ and use a Newton method to solve the algebraic system until the residual of the problem is less than 10^{-10} . We plot the solution for various values of α as well as a cross section through the coordinate axis. Notice that the solution becomes extremely badly behaved as α increases.

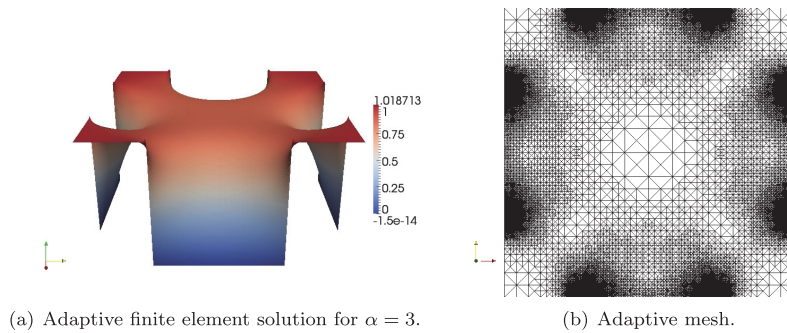


FIG. 6.3. Numerical results for a solution to Pucci's equation with a piecewise defined boundary condition (6.10). We choose $p = 2$ and use an adaptive scheme based on ZZ gradient recovery. The mesh is refined correctly about the jumps on the boundary.

7. Conclusions and outlook. In this work we have proposed a novel numerical scheme for fully nonlinear and generic quasi-linear PDEs. The scheme was based on a previous work for nonvariational PDEs (those given in nondivergence form) Lakkis and Pryer (2011a).

We have illustrated the application of the method for a simple, nonphysically motivated example, moving on to more interesting problems, that of the Monge–Ampère equation and Pucci’s equation.

For Pucci’s equation we numerically showed convergence and conducted experiments which may be compared with previous numerical studies.

We demonstrated for classical and singular viscosity solutions to the Monge–Ampère equation that the method is robust again, showing numerical convergence.

We postulate that the method is better suited to a discontinuous Galerkin framework, which is the subject of ongoing research.

REFERENCES

- N. E. AGUILERA AND P. MORIN (2009), *On convex functions and the finite element method*, SIAM J. Numer. Anal., 47, pp. 3139–3157.
- M. AINSWORTH AND J. T. ODEN (2000), *A posteriori error estimation in finite element analysis*, in Pure and Applied Mathematics, Wiley-Interscience, New York.
- G. AWANOU (2010), *Spline Element Method for the Monge–Ampère Equation*, preprint, Northern Illinois University, <http://www.math.niu.edu/~awanou/SplineMonge7.pdf>.
- G. AWANOU (2011), *Pseudo Time Continuation and Time Marching Methods for Monge–Ampère Type Equations*, preprint, <http://www.math.niu.edu/~awanou/MongePseudo05.pdf>.
- G. BARLES AND P. E. SOUGANIDIS (1991), *Convergence of approximation schemes for fully nonlinear second order equations*, Asymptot. Anal., 4, pp. 271–283.
- K. BÖHMER (2008), *On finite element methods for fully nonlinear elliptic equations of second order*, SIAM J. Numer. Anal., 46, pp. 1212–1249.
- S. C. BRENNER, T. GUDI, M. NEILAN, AND L.-Y. SUNG (2011), *C^0 penalty methods for the fully nonlinear Monge–Ampère equation*, Math. Comp., 80, pp. 1979–1995.
- L. A. CAFFARELLI AND X. CABRÉ (1995), *Fully Nonlinear Elliptic Equations*, Amer. Math. Soc. Colloq. Publ. 43, AMS, Providence, RI.
- P. G. CIARLET (1978), *The Finite Element Method for Elliptic Problems*, Stud. Math. Appl. 4, North-Holland, Amsterdam.
- O. DAVYDOV AND A. SAEED (2010), *Stable Splitting of Bivariate Splines Spaces by Bernstein–Bézier Methods*, in “Curves and Surfaces 7th International Conference, Avignon, France, June 24–30, 2010” (J.-D. Boissonnat et al, Eds.), pp. 220–235, LNCS 6920, Springer-Verlag, 2012.
- O. DAVYDOV AND A. SAEED (2012), *Numerical Solution of Fully Nonlinear Elliptic Equations by Böhmer’s Method*, J. Comput. Appl. Math., to appear.
- E. J. DEAN AND R. GLOWINSKI (2003), *Numerical solution of the two-dimensional elliptic Monge–Ampère equation with Dirichlet boundary conditions: An augmented Lagrangian approach*, C. R. Math. Acad. Sci. Paris, 336, pp. 779–784.
- E. J. DEAN AND R. GLOWINSKI (2005), *On the numerical solution of a two-dimensional Pucci’s equation with Dirichlet boundary conditions: A least-squares approach*, C. R. Math. Acad. Sci. Paris, 341, pp. 375–380.
- E. J. DEAN AND R. GLOWINSKI (2006), *Numerical methods for fully nonlinear elliptic equations of the Monge–Ampère type*, Comput. Methods Appl. Mech. Engrg., 195, pp. 1344–1386.
- L. C. EVANS (1998), *Partial Differential Equations*, Grad. Stud. Math., AMS, Providence, RI.
- X. FENG AND M. NEILAN (2009a), *Mixed finite element methods for the fully nonlinear Monge–Ampère equation based on the vanishing moment method*, SIAM J. Numer. Anal., 47, pp. 1226–1250.
- X. FENG AND M. NEILAN (2009b), *Vanishing moment method and moment solutions for fully nonlinear second order partial differential equations*, J. Sci. Comput., 38, pp. 74–98.
- B. D. FROESE (2011), *A Numerical Method for the Elliptic Monge–Ampère Equation with Transport Boundary Conditions*, Technical report, arXiv:1101.4981v1.
- M. JENSEN AND I. SMEARS (2011), *On the Convergence of Finite Element Methods for Hamilton–Jacobi–Bellman Equations*, Technical report, arXiv:1111.5423.
- C. T. KELLEY (1995), *Iterative Methods for Linear and Nonlinear Equations*, Frontiers in Appl. Math., SIAM, Philadelphia.

- H. J. KUO AND N. S. TRUDINGER (1992), *Discrete methods for fully nonlinear elliptic equations*, SIAM J. Numer. Anal., 29, pp. 123–135.
- H.-J. KUO AND N. S. TRUDINGER (2005), *Estimates for solutions of fully nonlinear discrete schemes*, in Trends in Partial Differential Equations of Mathematical Physics, Progr. Nonlinear Differential Equations Appl. 61, Birkhäuser, Basel, pp. 275–282.
- O. LAKKIS AND T. PRYER (2011a), *A finite element method for second order nonvariational elliptic problems*, SIAM J. Sci. Comput., 33, pp. 786–801.
- O. LAKKIS AND T. PRYER (2011b), *Gradient recovery in adaptive methods for parabolic problems*, IMA J. Numer. Anal., 32, pp. 246–278.
- G. LOEPER AND F. RAPETTI (2005), *Numerical solution of the Monge-Ampère equation by a Newton’s algorithm*, C. R. Math. Acad. Sci. Paris, 340, pp. 319–324.
- A. LOGG AND G. N. WELLS (2010), *DOLFIN: Automated finite element computing*, ACM Trans. Math. Software, 37.
- M. NEILAN (2012), *Finite Element Methods for Fully Nonlinear Second Order PDE’s Based on the Discrete Hessian*, preprint.
- A. M. OBERMAN (2008), *Wide stencil finite difference schemes for the elliptic Monge-Ampère equation and functions of the eigenvalues of the Hessian*, Discrete Contin. Dyn. Syst. Ser. B, 10, pp. 221–238.
- V. I. OLIKER AND L. D. PRUSSNER (1988), *On the numerical solution of the equation $(\partial^2 z / \partial x^2)(\partial^2 z / \partial y^2) - ((\partial^2 z / \partial x \partial y))^2 = f$ and its discretizations*, I. Numer. Math., 54, pp. 271–293.
- T. PRYER (2010), *Recovery Methods for Evolution and Nonlinear Problems*, D.Phil. thesis, University of Sussex; also available online from <http://sro.sussex.ac.uk/6285>.
- H. J. STETTER (1973), *Analysis of Discretization Methods for Ordinary Differential Equations*, Springer Tracts Natural Philosophy 23, Springer-Verlag, New York.